CEO Attributes, Compensation, and Firm Value: Evidence from a Structural Estimation

Internet Appendix

A. Participation constraint

In evaluating when the participation constraint binds, we consider three cases: $\gamma > 1$, $\gamma < 1$, and $\gamma = 1$ (log utility).

Proposition 1. *Define:*

$$\Delta = U^R\left(a_{t-1}, z_{t-1}, \delta e_{t-1}, t-1\right) - \beta E_{t-1}\left[U^R\left(a_t, z_t, e_t, t\right)\right] + E_{t-1}\left[k\left(a_t\right)\right].$$

Given a next period equity share of e_t, a) for $\gamma \geq 1$, there exists $f_t > 0$ at which the participation constraint binds when:

$$\Delta < \frac{1}{\gamma - 1};$$

b) for $\gamma < 1$, there exists $f_t \geq 0$ at which the participation constraint binds when:

$$\Delta \geq \frac{1}{\gamma - 1};$$

and c) for $\gamma = 1$, there always exists $f_t > 0$ at which the participation constraint binds.

Proof. Recall the participation constraint given in inequality (16):

$$E_{t-1}\left[U\left(z_t, f_t, e_t, t\right)\right] \geq U^R\left(a_{t-1}, z_{t-1}, \delta e_{t-1}, t-1\right).$$

Using the definition of the CEO's utility, and some algebra, we can rewrite the participation constraint as:

$$\frac{f_t^{1-\gamma} - 1}{1 - \gamma} \geq \Delta. \tag{A.1}$$

For part (a) of Proposition 1, rewrite (A.1) for $\gamma > 1$:

$$f_t \leq (1 - \gamma)\,\Delta + 1.$$

We want this inequality to hold for some $f_t > 0$, which is true if and only if:

$$\Delta < \frac{1}{\gamma - 1}.$$

For part (b) of Proposition 1, rewrite (A.1) for $\gamma < 1$:

$$f_t \geq (1-\gamma)\,\Delta + 1.$$

We want this inequality to hold for some $f_t \geq 0$, which is true if and only if:

$$\Delta \geq \frac{1}{\gamma - 1}.$$

For part (c) of Proposition 1, using part (a):

$$\lim_{\gamma \to 1^+} \frac{1}{\gamma - 1} = +\infty,$$

which implies $\Delta < \frac{1}{\gamma-1}$ always holds true for $\gamma = 1$. □

Note that, for $\gamma > 1$, $\Delta \geq \frac{1}{\gamma-1}$ implies that there is no wage large enough to induce the CEO to work for the firm another year. As a result, he retires and receives his retirement utility. Even when the participation constraint will not hold (i.e., the CEO knows the board cannot offer him a large enough wage to stay), the assumed form of the CEO's maximization problem in equation (5) is still the correct one.

For $\gamma < 1$, $\Delta < \frac{1}{\gamma-1}$ implies that no cash payments are necessary to induce the CEO to work for the firm. Negative salaries are not generally allowed, as the CEO's utility function does not have a real value for $f_t < 0$ for many values of $\gamma \in (0, 1)$.

B. Fit and robustness

In this section we examine the model's fit to the data, both dynamically and cross-sectionally. Then we examine the robustness of the model to three assumptions. First, we estimate the model excluding CEO influence. Next, we estimate the model using different discount factors. Last, we look at whether the model's estimates are stable over time.

B.1. Model fit: Dynamic

Table A1 presents regressions of logged contracting variables (equity ownership, cash salary and total compensation) on lagged ownership, salary, compensation, firm value and CEO tenure, both for the data and the model simulation, using the full sample estimation.

Panel A presents the actual and simulated regressions excluding the lagged dependent variable, while Panel B includes the lagged dependent variable. These dynamic regressions serve as an external check on the model's performance, given that the model is not estimated to specifically match these statistics. This comparison is not a statistical test, and, as such, is more qualitative in nature.

The equity ownership dynamics are quite similar for both the model and the data. When excluding the lag of equity ownership, all of the signs are the same, and the magnitudes are similar. This similarity mostly goes through when the lag is included, with the main difference being the sign on firm value no longer matching. While lagged firm value does enter this regression significantly, its economic importance is quite low. In a standardized regression it has practically zero effect (0.024), with only the coefficient on tenure being smaller. The conclusion from the equity regressions is that the model is able to replicate equity ownership dynamics well.

The model is not as successful duplicating the salary and total compensation dynamics. In most respects, especially when the lag of the dependent value is included, the simulated dynamics compare favorably to the data dynamics. The model primarily struggles to estimate the sensitivity of salary to equity ownership, and of total compensation to firm value. In both cases, the model produces sensitivities that are too large. The sensitivity of total compensation to firm value is explained by firm value being less volatile in the simulation than in the data; a standardized version of this same regression produces estimates that are of a similar magnitude (0.428 in the simulation vs. 0.286 in the data). Thus it is only the large sensitivity of salary to equity ownership that is of significant concern. The model is relatively simple, and does not include any dynamics concerning the CEO's outside opportunity.[17] Such dynamics would likely break the close tie of equity ownership and salary, allowing it to more successfully replicate salary dynamics.

[17] For example, shocks to the CEO's outside wealth, changes in the CEO's job market or learning about the CEO's abilities as in Taylor (2013).

Given the model's success in replicating the level of equity ownership, salary and total compensation, along with its ability to replicate equity ownership dynamics and some salary and total compensation dynamics, the model appears to be a useful simplification of CEO contracting.

B.2. Model fit: Cross-sectional

We look at the model's fit in the cross-section in two ways. The first is comparing the cross-section of average compensation variables in the data and in the simulations. The second is looking at the partial covariance of contracting variables along with firm size and CEO tenure in the cross-section. As in Section B.1, the comparisons are qualitative.

Figure A1 compares the mean values of the log of CEO ownership, the log of salary and the log of total compensation for the data and the simulations of the 54 representative firms. Most of the estimated ownership and salary values, which are used in estimation, sit near the forty-five degree line, indicating the model is flexible enough to fit these moments in the cross section. Total compensation is not used for estimation, and the resulting match is not as strong, with many of the values sitting above the forty-five degree line. Even though the total compensation values do not match as well, they still plot along the line, indicating the model does capture the direction of their distribution. Overall, the model is able to replicate the distribution of compensation in the cross section.

Table A2 presents the regression of CEO ownership, salary and total compensation on each other plus firm value for the actual data and simulated data. Comparing the simulations with the data, the signs, magnitudes and significance levels are practically the same for all three regressions. These results show that the model is able to replicate how compensation variables covary between firms. This finding suggests the model is flexible enough to match a range of contracting environments, and is of general use in understanding cross-sectional compensation variation.

B.3. No influence estimation

Table A3 presents the model estimation on the full sample when CEO influence is required to be zero. The model's fit is substantially worse than when we include CEO influence. As we might expect, moments related to compensation are the most difficult to match without influence. Of note, the model is incapable of matching the level of CEO equity exposure without including CEO influence, which supports the results above that equity ownership is higher than necessary for the typical firm, as the model seems unable to generate sufficient equity exposure using risk aversion, effort aversion and reservation utility alone.

B.4. Discount factor

For the estimation, we assume that the shareholder's, and CEO's, discount factor reflects their expected return on equity. As the sample mean return is 7.76%, we use a discount factor of 0.928. This value is calibrated to data in advance of the primary parameter estimation, so there could be concern as to how it affects the parameter estimates. To get a sense for its effect, we estimate the model on the full sample using two different discount factors: 0.875 and 0.975. The parameter estimates are presented in Panel A of Table A4. When combined with the full sample results using a discount factor of 0.928, there does appear to be a trend in most of the parameters estimates. Risk aversion, CEO influence, cash flow volatility and firm scale all decrease as the discount rate increases. Reservation value and cash flow autocorrelation increase. Effort aversion does not seem to have a trend.

As we might expect, parameter estimates are sensitive to the assumed discount factor. There is a connection between time preference and risk aversion that the model ignores by not estimating the discount factor. As time preferences increase in importance, proxied by a larger discount factor, there is a decrease in estimated risk aversion. As a result of lower risk aversion, less CEO influence is necessary to explain the CEO delta level. With less influence, a larger reservation value is necessary to match the cash compensation level. Additionally, with lower risk aversion the mean effort level increases, which allows the model to work with a lower estimate of scale.

Changing the discount factor also affects the estimate of the benefit from removing CEO influence. Qualitatively everything is the same: shareholders gain from removing CEO influence, the average salary is slightly higher, and total compensation and delta are strongly decreased. Quantitatively, there are differences. For example, for a discount factor of 0.875, removing CEO influence increases shareholder value by 4.5%. The increase is 0.4% using a discount factor of 0.975.

Overall, which discount factor we use matters for the model's estimates, and the magnitude of effect of removing influence. In comparing the model fit between the different discount factor assumptions, we find that the original estimation has the best fit.

B.5. Time periods

The data spans the entire 1992 to 2014 time frame, and the estimated parameters are for the typical firm in the sample. However, we know that the sample has undergone a number of changes over the twenty-two years. For example, the tech boom and subsequent bust of the late 1990s changed the composition of the sample. In terms of data quality, FAS 123r changed the reporting of stock option grants, meaning the delta computations post 2006 are higher quality than those from before. Additionally, as noted by Bettis, Bizjak, Coles, and Kalpathy (2010), the nature of how options and restricted stock vest has changed over the sample, shifting from being time based to being more performance based. To understand how these changes affect the model's estimates, we split the sample into three time periods: early (before 2001), late (after 2006) and middle (in between). Panel B of Table A4 reports the parameter estimates.

The parameter estimates are quite stable over time. The early sample is the most different of the three, with lower risk aversion, reservation value and CEO influence. Still, they are similar to each other, and to the full sample parameter estimates. The effect of removing influence is quite similar between the three and the full sample as well. Removing influence increases shareholder value by 0.7%, 1.4% and 1.5% for the early, middle and late time periods, respectively. The compensation effects from removing influence are also of a similar magnitude with the full sample values. These similarities across time give us confidence that changes in the sample do not significantly bias our results.

C. Comparative Statics

To produce comparative statics, we use the parameter estimates from the main estimation as the baseline. Next, one at a time, we change a parameter, re-solve and re-simulate the model, then calculate the statistic using the new simulation. We then plot these statistics as functions of the parameter value.

Figure A2 presents comparative statics for the compensation variables (delta, cash compensation and total compensation) in relation to the four CEO attribute parameters: risk aversion, γ, effort aversion, ψ, reservation value, ω, and influence, λ. Note that, to give a sense of the cross-sectional variation in the attributes, the shaded area of each figure corresponds to the interquartile range of the estimates, given in Table 4, and the vertical dotted line is the value found in the full sample estimation. Increasing risk aversion decreases total pay and equity exposure, with little effect on cash compensation. Given their role in the participation constraint, both effort aversion and the reservation value have strong positive effects on all components of pay, although total pay and cash pay both flatten out at high levels of each as boards transition to greater amounts of equity-based pay. Influence has a positive effect on delta and total pay, with the greatest effect on delta. Figure A3 presents comparative statics for firm and shareholder value. Both firm value and shareholder value decrease as all of the parameter values increase. The largest effects on shareholder value come from increasing effort aversion and reservation value.

We here provide more detail on the two "surprising" comparative statics discussed in section 3.1: the positive relation between effort aversion and delta, and the negative direction between firm scale and delta.

In Holmstrom (1979), an increase in the effort cost parameter decreases equity exposure, while effort aversion has the opposite effect in Table 6. The direction for this comparative static differs between the models because equity affects the CEO's effort decision differently.

Equity and effort are linearly related in Holmstrom (1979), so when effort becomes more costly, decreasing the slope of the relation, effort decreases more at higher equity levels. As Figure 1 shows, effort and equity are nonlinearly related in our model. This nonlinear relation is the result of our assumed effort cost function, which provides an upper bound on effort by forcing the cost to negative infinity as effort approaches one. More equity thus has a decreasing marginal effect on effort, as there is a limit to how much extra effort can be induced. In Holmstrom (1979), equity's marginal effect on effort is constant, as there is no similar upper bound. The largest decreases in effort for an increase in equity cost come at low levels of equity, where the slope of the equity-effort curve is steepest. In many simulated states, an increase in effort cost is optimally met with an increase in equity exposure, as the resulting decrease in CEO effort is smaller with more exposure. As a result, an increase in effort cost leads to an average increase in CEO delta in the simulations.

The negative relation between delta and firm scale is an artifact of how we model influence. All else equal, and ignoring any change to CEO effort, log firm value increases approximately linearly with scale, with a slope close to one. CEO utility does not increase at the same rate with an increase in scale, because of its curvature. As a result, for a given λ, an increase in η decreases the CEO's effective influence, as the shareholder value portion of the board of directors' problem increases by more than the CEO utility component. This decrease in effective influence leads to lower delta, as the CEO influences his compensation by less. In the absence of CEO influence, this comparative static is positive. This change in effective influence, despite a constant λ, is a reminder for us to be cautious in interpreting differences in the parameter value between firms of differing sizes; a more useful comparison is the change in compensation and firm value when we remove influence.

Table A1: Model Fit: Dynamic

This table compares dynamic regressions of compensation variables on lagged compensation, firm value and tenure run on data from the full sample and from a simulation of the estimateed model. *LogDelta* is defined as the log of one plus delta, *LogSalary* is defined as the log of one plus annual cash compensation, *LogPay* is defined as the log of one plus annual total compensation, *LogFirmValue* is defined as the log of market value, and *LogTenure* is the log of number of years that have elapsed since the CEO was hired. A *Lag* prefix denotes the value from the previous year. Panel A presents the regressions excluding the lagged dependent variable, and Panel B presents regressions including the lagged dependent variable. All regressions use fixed effects for firm and year to control for heterogeneity. Heteroscadisticity-robust standard errors are in parentheses. Stars by an estimate signals statistical significance: (***) for significance at the 1% level, (**) for 5% level, and (*) for 10% level.

Panel A. Without Lagged Dependent Variable

	Actual			Simulated		
	LogDelta	*LogSalary*	*LogPay*	*LogDelta*	*LogSalary*	*LogPay*
LagLogDelta		0.008**	0.016***		0.907***	0.605***
		(0.004)	(0.006)		(0.006)	(0.009)
LagLogSalary	0.374***		0.401***	0.295***		−0.086***
	(0.022)		(0.013)	(0.005)		(0.004)
LagLogPay	0.313***	0.197***		0.168***	0.009***	
	(0.015)	(0.007)		(0.002)	(0.001)	
LagLogFirmValue	0.389***	0.084***	0.223***	0.385***	−0.037***	0.829***
	(0.009)	(0.004)	(0.006)	(0.006)	(0.006)	(0.011)
LogTenure	0.410***	0.054***	−0.011*	0.223***	0.076***	0.063***
	(0.010)	(0.004)	(0.006)	(0.003)	(0.003)	(0.005)
R^2	0.330	0.166	0.214	0.836	0.945	0.757

Panel B. With Lagged Dependent Variable

	Actual			Simulated		
	LogDelta	*LogSalary*	*LogPay*	*LogDelta*	*LogSalary*	*LogPay*
LagLogDelta	0.817***	−0.019**	0.003	0.817***	0.899***	0.240***
	(0.008)	(0.004)	(0.006)	(0.007)	(0.007)	(0.011)
LagLogSalary	0.076***	0.414***	0.296***	0.022***	0.011***	0.177***
	(0.012)	(0.014)	(0.019)	(0.004)	(0.005)	(0.006)
LagLogPay	0.063***	0.012	0.120***	0.030***	0.012***	0.167***
	(0.009)	(0.010)	(0.014)	(0.002)	(0.002)	(0.003)
LagLogMarketValue	0.032***	0.109***	0.218***	−0.074***	−0.037***	0.918***
	(0.007)	(0.004)	(0.006)	(0.006)	(0.006)	(0.010)
LogTenure	−0.011*	0.040***	0.001	0.019***	0.073***	0.079***
	(0.006)	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)
R^2	0.764	0.292	0.224	0.895	0.945	0.794

Table A2: Model Fit: Cross Section

This table compares cross-sectional regressions of compensation variables on other compensation variables and firm value for the cross section of 54 representative firms in the actual data and the simulated data. *LogDelta* is defined as the log of one plus delta, *LogSalary* is defined as the log of one plus annual cash compensation, *LogPay* is defined as the log of one plus annual total compensation, and *LogFirmValue* is defined as the log of market value. Each value is the mean value for that variable in each of the 54 subsamples. Heteroscadisticity-robust standard errors are in parentheses. Stars by an estimate signals statistical significance: (***) for significance at the 1% level, (**) for 5% level, and (*) for 10% level.

	Actual			Simulated		
	LogDelta	*LogSalary*	*LogPay*	*LogDelta*	*LogSalary*	*LogPay*
LogDelta		−0.087***	0.107***		−0.123***	0.143***
		(0.029)	(0.039)		(0.034)	(0.039)
LogSalary	−2.141***		0.956***	−2.359***		0.985***
	(0.590)		(0.137)	(0.634)		(0.124)
LogPay	1.893**	0.684***		1.984***	0.713***	
	(0.871)	(0.057)		(0.719)	(0.053)	
LogFirmValue	0.362	0.033	0.073*	0.309*	0.024	0.074**
	(0.232)	(0.022)	(0.042)	(0.172)	(0.022)	(0.033)
R^2	0.770	0.936	0.960	0.815	0.944	0.968

Table A3: Parameter Estimation: Full Sample without CEO Influence

This table reports estimation results based on the full sample, but removing CEO influence on the board from the model. The estimation is done using a simulated minimum distance estimator, which matches moments from model simulations to the same moments in the data to identify the structural model parameters. Panel A reports the simulated and actual moments, along with firm-level clustered t-statistics for the differences between corresponding moments. Note that log delta (salary) is actually the natural logarithm of one plus delta (salary), which allows for zero values of delta (salary). Sensitivity is computed as the parameter from a regression of the log of one plus delta (salary) on the log of firm value. The sensitivity, serial correlation, and variance moments are all computed using within-firm variation alone. Panel B reports the the estimates and firm-level clustered standard errors (in parentheses) of the model's parameters: CEO risk aversion γ, CEO effort aversion ψ, CEO reservation value ω, CEO influence with the board λ, the autocorrelation of the cash flow law of motion ρ, the standard deviation of cash flow shocks σ, and cash flow scale η.

Panel A. Moments

	Actual Moments	Simulated Moments	T-statistics
Mean log delta	3.228	3.083	−4.701
Mean log salary	1.086	1.175	10.929
Mean salary to total compensation ratio	47.301	48.355	6.368
Mean log firm value	7.690	7.689	−0.012
Sensitivity of delta to firm value	0.759	0.907	42.504
Sensitivity of salary to firm value	0.190	0.234	44.441
Serial correlation of log delta	0.859	0.840	−3.603
Serial correlation of log salary	0.424	0.695	22.875
Serial correlation of cash flow	0.853	0.756	−7.531
Variance of cash flow growth	0.211	0.174	−7.543

Panel B. Parameters

γ	ψ	ω	λ	ρ	σ	η
1.282	0.523	1.903	0.000	0.818	0.465	5.541
(0.044)	(0.014)	(0.057)	n/a	(0.006)	(0.007)	(0.004)

Table A4: Parameter Estimation: Robustness

This table reports estimation results from using a different assumed discount rate, and from splitting the sample by time period. The estimation is done using a simulated minimum distance estimator, which matches moments from model simulations to the same moments in the data to identify the structural model parameters. Panel A shows parameter estimates when the discount factor, β is varied and Panel B shows the parameter estimates for the early (before 2001), middle (2001-2006) and late (after 2006) periods. The model's parameters are: CEO risk aversion γ, CEO effort aversion ψ, CEO reservation value ω, CEO influence with the board λ, the autocorrelation of the cash flow law of motion ρ, the standard deviation of cash flow shocks σ, and cash flow scale η. Firm-level clustered standard errors are in parentheses.

γ	ψ	ω	λ	ρ	σ	η
Panel A. Discount factor						
$\beta = 0.875$						
2.495	0.271	2.703	0.607	0.810	0.539	6.190
(0.008)	(0.004)	(0.014)	(0.002)	(0.007)	(0.007)	(0.008)
$\beta = 0.975$						
1.809	0.228	3.372	0.133	0.923	0.443	4.509
(0.032)	(0.003)	(0.035)	(0.025)	(0.001)	(0.005)	(0.005)

Panel B. Time periods						
Early						
1.706	0.317	2.616	0.206	0.832	0.462	5.385
(0.014)	(0.003)	(0.014)	(0.038)	(0.006)	(0.011)	(0.039)
Middle						
2.037	0.238	3.195	0.499	0.812	0.493	5.563
(0.026)	(0.007)	(0.023)	(0.037)	(0.018)	(0.013)	(0.035)
Late						
2.070	0.309	2.893	0.343	0.825	0.518	5.760
(0.014)	(0.006)	(0.009)	(0.018)	(0.005)	(0.011)	(0.016)

Figure A1: Model Fit: Cross Section

Figure A1 shows scatter plots of actual (x-axis) and simulated (y-axis) mean log of one plus CEO delta, *LogDelta*, log of one plus cash compensation, *LogSalary*, and log of one plus total compensation, *LogPay*, for the 54 representative firms cross section. The solid line is the 45° line.

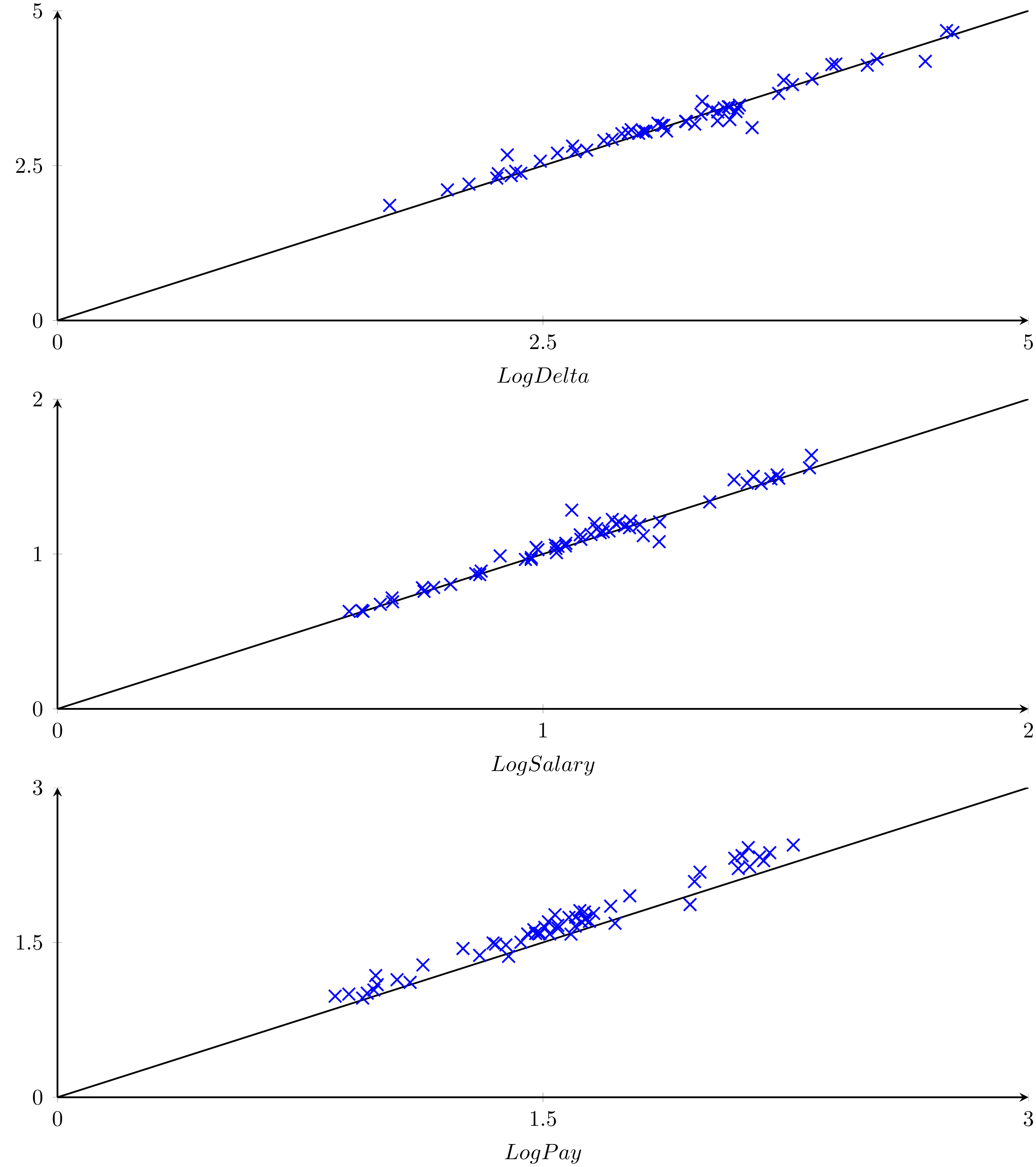

Figure A2: Comparative Statics: Compensation

Figure A2 shows how the mean simulated CEO equity exposure, *Delta*, cash compensation, *Salary* and total compensation, *Pay* change as CEO risk aversion (γ), effort aversion (ψ), reservation value (ω) and influence (λ) change. The vertical dotted line identifies the estimated value for each parameter from the full sample, while the gray area indicates the interquartile range of parameter estimates from the 54 representative firm cross section. Note that the y-axis is logarithmic.

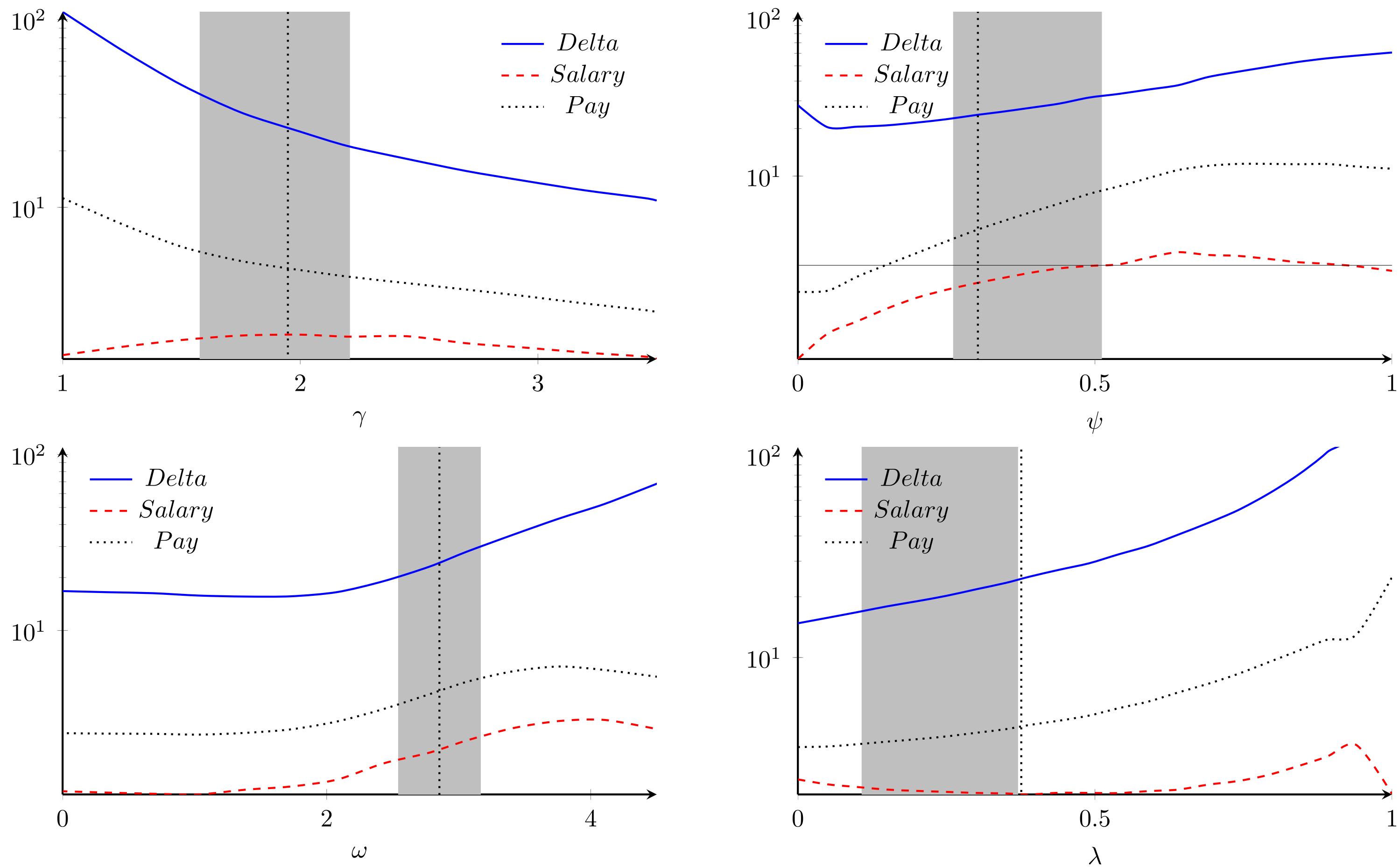

Figure A3: Comparative Statics: Value

Figure A3 shows how the mean simulated firm value and shareholder value change as CEO risk aversion (γ), effort aversion (ψ), reservation value (ω) and influence (λ) change. The vertical dotted line identifies the estimated value for each parameter from the full sample, while the gray area indicates the interquartile range of parameter estimates from the 54 representative firm cross section.